

## Modelo de predição de risco de internação em pacientes com Diabetes *mellitus* tipo 2 atendidos na atenção primária de Curitiba – uma abordagem da inteligência artificial com machine learning

Guilherme Silva Julian, Daniel Ferraz de Campos Filho

IQVIA Brasil

**Introdução:** O diabetes mellitus tipo 2 (DM2) é a doença mais prevalente mundialmente. Os agravamentos da DM2 geram um ônus econômico de até US\$ 827 bilhões anuais mundialmente. Dentre estes gastos, estima-se que a maior parte é relacionada à medicamentos e internações hospitalares. No Brasil, existem inequidades que impactam em acesso ao sistema de saúde e também em desfechos. Especificamente em DM2, o cuidado ao paciente é principalmente realizado na atenção primária. Por outro lado, a previsibilidade do perfil de pacientes com DM2 que potencialmente serão internados é importante para um melhor manejo da doença com intervenções e gestão de saúde, com planejamento de ocupação de leitos. **Objetivos:** Identificar, descrever e construir um modelo de predição de risco de internação em pacientes de DM2 atendidos nas unidades de atenção primária da cidade de Curitiba. **Métodos:** Extraímos dados de pacientes com CID-10 E11 registrados de novembro de 2016 até maio de 2019 nos bancos de dados de consultas com profissionais de saúde em unidades de atenção primária da cidade de Curitiba. Apenas registros únicos foram considerados. Identificamos pacientes únicos com ao menos um registro de CID-10 E11 e avaliamos longitudinalmente estes pacientes. Em seguida, desenvolvemos um modelo de predição a partir de 38 variáveis, das quais 10 foram incluídas como preditoras. Foram treinados 8 algoritmos de machine learning. Após treinados, os mesmos foram submetidos à testes de acurácia, especificidade e sensibilidade entre si. Baseado nos **Resultados:** probabilísticos do modelo selecionado, grupos de risco para internação foram definidos. **Resultados:** No período avaliado, foram identificados 14.579.665 de registros de atendimento em atenção primária em um total de 1.072.986 pacientes. Destes pacientes, 24.058 tinham ao menos um registro de E11 e foram incluídos no modelo, sendo a maioria mulheres (58%) com idade mediana de 61 anos. A mediana de acurácia dos 8 modelos treinados foram de 0.67 a 0.71. Dos modelos testados, Stochastic Gradient Boosting, Neural Network e Extreme Gradient Boosting foram selecionados para a etapa comparativa de teste na base completa, sendo XGB o modelo selecionado, com melhor performance e área sob a curva ROC de 0.81. Com os **Resultados:** do modelo, classificamos os pacientes em baixo, médio e alto risco de internação. No modelo final, 2.034 pacientes foram classificados como de alto risco, 8.202 como médio risco e 13.822 como baixo risco. Dentre os pacientes de alto risco, 1.364 pacientes não tiveram internação no período avaliado. **Conclusão:** A utilização de modelos preditivos pode auxiliar na identificação de pacientes de alto risco, antecipando uma potencial internação hospitalar dos pacientes, o que permite realizar intervenções – medicamentosas ou não – que melhorem o manejo destes pacientes. O modelo proposto foi capaz de identificar 1.364 pacientes de alto risco de internação que podem ter o evento evitado com intervenções.